PATENT ABSTRACTS OF JAPAN

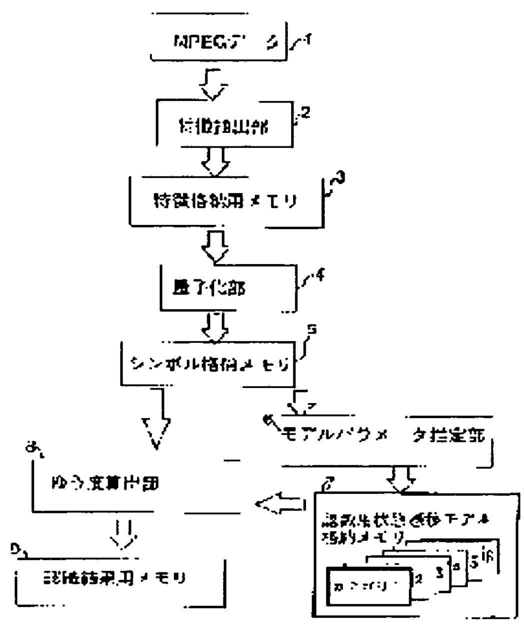
(11)Publication number : 10-013832
(43)Date of publication of application : 16.01.1998

(51)Int.Cl. H04N 7/30
G06T 7/20

(21)Application number : 08-164430
(22)Date of filing : 25.06.1996

(71)Applicant : NIPPON TELEGR & TELEPH CORP <NTT>
(72)Inventor : YAMATO JUNJI
MURASE HIROSHI

(54) MOVING PICTURE RECOGNIZING METHOD AND MOVING PICTURE RECOGNIZING AND RETRIEVING METHOD



(57)Abstract:
PROBLEM TO BE SOLVED: To directly recognize and retrieve a specific pattern from compressed moving picture data through the use of the parameter of maximum likelihood with respect to a symbol string by extracting a DCT coefficient from image data and transforming its feature vector string to the symbol string.
SOLUTION: One frame of image data is divided by MPEG data 1 to obtain the DCT(discrete cosine transformation) coefficient of each unit. Next, a feature extraction part 2 fetches the frame feature vector of a low frequency component. Similarly, the feature vector string of a series of moving picture data is fetched and recorded in a memory for storing feature 3. Next, a quantization part 4 vector-quantizes it and records the symbol string in a symbol storing memory 5. A model parameter estimating part 6 estimates the parameter of such a state transition model as generates this symbol string and records it in a state transmission model storing memory for recognition 7. A likelihood calculating part 8 estimates the parameter of the model of high likelihood for each category of a recognizing object and stores it in a memory for a recognizing result 9. Thereby the specific moving picture pattern is recognized and retrieved from compressed moving picture data.

LEGAL STATUS
[Date of request for examination]
[Date of sending the examiner's decision of rejection]
[Kind of final disposal of application other

than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

CLAIMS

[Claim(s)]
[Claim 1] In the dynamic-image recognition approach of recognizing the dynamic-image pattern of a series of dynamic images The step which extracts a break and the DCT multiplier of each block for the image data of each screen which displays a series of dynamic images to the block of MxN, By the feature-vector train of the step which extracts at least one of said the DCT multipliers as a feature vector of each screen, and the time series which consists of feature vectors of each screen which displays a specific dynamic-image pattern The step which learns a probable state-transition model for every dynamic-image pattern of two or more specification with which it becomes a recognition key, The feature-vector train of the time series which consists of feature vectors extracted from the image data of each screen which displays a series of dynamic images which are the candidates for recognition, The dynamic-image recognition approach characterized by providing the step which outputs the dynamic-image pattern of the state-transition model with which the likelihood to two or more state-transition models obtained by said study serves as max as a recognition result.
[Claim 2] The dynamic-image recognition approach indicated by claim 1 characterized by the image data of each screen which displays a series of dynamic images which are said candidates for recognition using a part of DCT multiplier contained in the image data of each screen which is compressed by the standard coding method and compressed by the standard coding method as a feature vector of each screen.
[Claim 3] The dynamic-image recognition approach indicated by claim 1 characterized by using a motion vector with a DCT multiplier as a feature vector of each of said screen.
[Claim 4] The dynamic-image recognition approach indicated by claim 3 to which it is characterized by using the part and motion compensation vector of the DCT multiplier contained in the image data of each screen which the image data of each screen which displays a series of dynamic images which are said candidates for recognition is compressed by the standard coding method, and was compressed by the standard coding method as a feature vector of each screen.
[Claim 5] The image recognition approach indicated by claim 2 or claim 4 characterized by using it among the DCT multipliers contained in the image data of each screen compressed by said standard coding method, carrying out the feature vector of 3 thru/or the DCT multiplier of 21 low-frequency components.
[Claim 6] The image recognition approach indicated by claim 2 or claim 4 characterized by using it among the DCT multipliers contained in the image data of each screen compressed by said standard coding method, carrying out the feature vector of the DCT multiplier on the 1st horizontal Rhine.

[Claim 7] The image recognition approach indicated by claim 2 or claim 4 characterized by using it among the DCT multipliers contained in the image data of each screen compressed by said standard coding method, carrying out the feature vector of the DCT multiplier 1st on [of the perpendicular approach] Rhine.
[Claim 8] The image recognition approach indicated by claim 2 or claim 4 characterized by using it, carrying out the feature vector of the DCT multiplier on the diagonal line which contains a dc component among the DCT multipliers contained in the image data of each screen compressed by said standard coding method.
[Claim 9] In the dynamic-image recognition search method which extracts the time domain containing a specific dynamic-image pattern out of a series of dynamic images The step which extracts a break and the DCT multiplier of each block for the image data of each screen which displays a series of dynamic images to the block of MxN, By the feature-vector train of the time series which consists of feature vectors of each screen which displays the step which extracts at least one of said the DCT multipliers as a feature vector of each screen, and the specific dynamic-image pattern used as a search key In the feature-vector train of the time series which consists of a step which learns a probable state-transition model, and a feature vector extracted from the image data of each screen which displays a series of dynamic images which are the candidates for retrieval The dynamic-image recognition search method characterized by providing the step which outputs the time domain where the likelihood to the state-transition model obtained by said study is high as a retrieval result.
[Claim 10] The dynamic-image recognition search method indicated by claim 9 characterized by the image data of each screen which displays a series of dynamic images which are said candidates for retrieval using a part of DCT multiplier contained in the image data of each screen which is compressed by the standard coding method and compressed by the standard coding method as a feature vector of each screen.
[Claim 11] The dynamic-image recognition search method indicated by claim 9 characterized by using a motion vector with a DCT multiplier as a feature vector of each of said screen.
[Claim 12] The dynamic-image recognition search method indicated by claim 11 to which it is characterized by using the part and motion compensation vector of the DCT multiplier contained in the image data of each screen which the image data of each screen which displays a series of dynamic images which are said candidates for retrieval is compressed by the standard coding method, and was compressed by the standard coding method as a feature vector of each screen.
[Claim 13] The image recognition search method indicated by claim 10 or claim 12 characterized by using it among the DCT multipliers contained in the image data of each screen compressed by said standard coding method, carrying out the feature vector of 3 thru/or the DCT multiplier of 21 low-frequency components.
[Claim 14] The image recognition search method indicated by claim 10 or claim 12 characterized by using it among the DCT multipliers contained in the image data of each screen compressed by said standard coding method, carrying out the feature vector of the DCT multiplier on the 1st horizontal Rhine.
[Claim 15] The image recognition search method indicated by claim 10 or claim 12 characterized by using it among the DCT multipliers contained in the image data of each screen compressed by said standard coding method, carrying out the feature vector of the DCT multiplier 1st on [of the perpendicular approach] Rhine.
[Claim 16] The image recognition search method indicated by claim 10 or claim 12 characterized by using it, carrying out the feature vector of the DCT multiplier on the

diagonal line which contains a dc component among the DCT multipliers contained in the image data of each screen compressed by said standard coding method.

[Translation done.]

DETAILED DESCRIPTION

[Detailed Description of the Invention]
[0001]
[Field of the Invention] This invention relates a specific dynamic-image pattern to the dynamic-image recognition approach and dynamic-image recognition search method which perform recognition and retrieval with respect to the dynamic-image recognition approach and a dynamic-image recognition search method out of the image data of each screen which displays a series of dynamic images especially.
[0002]
[Description of the Prior Art] Research of recent years many is done and the pattern recognition technique for a dynamic image has the well-known technique indicated by the following official report (b) as one of them.
[0003] (b) The technique of recognizing each actuation of animal objects, such as human being, is indicated by JP,5-46583,A aforementioned official report (**) (JP,5-46583,A) by symbol-izing the mesh description of the animal object extracted from the image data of each screen which displays a dynamic image by vector quantization, changing a dynamic-image sequence into a symbol sequence, and learning and recognizing the symbol sequence concerned.
[0004] Moreover, MEPG (Moving Picture Experts Group; international standards of media integrated system dynamic-image compression) and an international-standards coding method called MEPG2 are spreading as are recording of the dynamic-image data which constitute the nucleus technique of multimedia, or an information-compression technique in the case of transmission.
[0005]
[Problem(s) to be Solved by the Invention] Like the technique indicated by said official report (**) (JP,5-46583,A), when the dynamic-image pattern itself was searched for a specific dynamic-image pattern as a search key, mass image data and characteristic quantity data needed to be dealt with, and there was a trouble that the processing time of data processing increased out of the dynamic image of a conventional single string.
[0006] Moreover, when the standard coding method of MEPG and MEPG2 grade is spreading and it searches that dynamic-image pattern itself for a specific dynamic-image pattern as a search key out of a series of dynamic images, shortening the processing time of data processing is expected by using the dynamic-image data compressed by this standard coding method.
[0007] However, the optimal technique of searching a specific dynamic-image pattern out of a series of dynamic images for the dynamic-image data compressed by the standard coding method was not examined at all conventionally.

[0008] It is made in order that this invention may solve said trouble, and the purpose of this invention uses the dynamic-image data compressed by the standard coding method etc. in the dynamic-image recognition approach, and it is in offering the technique which becomes possible [shortening data-processing time amount].
[0009] In a dynamic-image recognition search method, other purposes of this invention use the dynamic-image data compressed by the standard coding method etc., and are to offer the technique which becomes possible [shortening data-processing time amount].
[0010] Other purposes and new descriptions are clarified by a publication and accompanying drawing of this specification at said purpose list of this invention.
[0011]
[Means for Solving the Problem] It will be as follows if the outline of a typical thing is briefly explained among invention indicated in this application.
[0012] (1) In the dynamic-image recognition approach of recognizing the dynamic-image pattern of a series of dynamic images The step which extracts a break and the DCT multiplier of each block for the image data of each screen which displays a series of dynamic images to the block of MxN, By the feature-vector train of the step which extracts at least one of said the DCT multipliers as a feature vector of each screen, and the time series which consists of feature vectors of each screen which displays a specific dynamic-image pattern The step which learns a probable state-transition model for every dynamic-image pattern of two or more specification with which it becomes a recognition key, The feature-vector train of the time series which consists of feature vectors extracted from the image data of each screen which displays a series of dynamic images which are the candidates for recognition, It is characterized by providing the step which outputs the dynamic-image pattern of the state-transition model with which the likelihood to two or more state-transition models obtained by said study serves as max as a recognition result.
[0013] (2) In the means of the above (1), the image data of each screen which displays a series of dynamic images which are said candidates for recognition is characterized by using a part of DCT multiplier contained in the image data of each screen which is compressed by the standard coding method and compressed by the standard coding method as a feature vector of each screen.
[0014] (3) In the means of the above (1), it is characterized by using a motion vector with a DCT multiplier as a feature vector of each of said screen.
[0015] (4) In the means of the above (3), the image data of each screen which displays a series of dynamic images which are said candidates for recognition is characterized by using the part and motion compensation vector of the DCT multiplier contained in the image data of each screen which is compressed by the standard coding method and compressed by the standard coding method as a feature vector of each screen.
[0016] (5) In the dynamic-image recognition search method which extracts the time domain containing a specific dynamic-image pattern out of a series of dynamic images The step which extracts a break and the DCT multiplier of each block for the image data of each screen which displays a series of dynamic images to the block of MxN, By the feature-vector train of the time series which consists of feature vectors of each screen which displays the step which extracts at least one of said the DCT multipliers as a feature vector of each screen, and the specific dynamic-image pattern used as a search key In the feature-vector train of the time series which consists of a step which learns a probable state-transition model, and a feature vector extracted from the image data of each screen which displays a series of dynamic images which are the

candidates for retrieval It is characterized by providing the step which outputs the time domain where the likelihood to the state-transition model obtained by said study is high as a retrieval result.

[0017] (6) In the means of the above (5), the image data of each screen which displays a series of dynamic images which are said candidates for retrieval is characterized by using a part of DCT multiplier contained in the image data of each screen which is compressed by the standard coding method and compressed by the standard coding method as a feature vector of each screen.

[0018] (7) In the means of the above (5), it is characterized by using a motion vector with a DCT multiplier as a feature vector of each of said screen.

[0019] (8) In the means of the above (7), the image data of each screen which displays a series of dynamic images which are said candidates for retrieval is characterized by using the part and motion compensation vector of the DCT multiplier contained in the image data of each screen which is compressed by the standard coding method and compressed by the standard coding method as a feature vector of each screen.

[0020] According to said each means, a DCT multiplier or a DCT multiplier, and a motion compensation vector are used as characteristic quantity, and since it has direct-recognized and the specific dynamic-image pattern was searched, it becomes possible to lessen the processing time of data processing from the dynamic-image data of the few capacity compressed by the standard coding method of MEPG and MEPG2 grade.

[0021]

[Embodiment of the Invention] Hereafter, the gestalt of implementation of invention of this invention is explained to a detail with reference to a drawing.

[0022] In addition, in the complete diagram for explaining the gestalt of implementation of invention, what has the same function attaches the same sign, and explanation of the repeat is omitted.

[0023] Drawing 1 is the functional block diagram showing the outline configuration of the dynamic-image recognition retrieval equipment with which the dynamic-image recognition approach and dynamic-image recognition search method which are the gestalt of implementation of 1 invention of this invention are applied.

[0024] drawing 1 -- setting -- 1 -- MEPG data and 2 -- the feature-extraction section and 3 -- for symbol storing memory and 6, as for the state-transition model storing memory for recognition, and 8, the model parameter estimation section and 7 are [the memory for the description storing, and 4 / the quantization section and 5 / the likelihood calculation section and 9] the memory for recognition results.

[0025] Here, as said state-transition model storing memory 7 for recognition, and memory 9 for recognition results, external storage is used and said MEPG data 1 are stored in external storage, for example.

[0026] There are three phases of study and recognition in fundamental actuation of the gestalt of operation of this invention, and at the time of study, parameter estimation of the state-transition model for recognition is performed from the data for study, and it stores in the state-transition model storing memory 7 for recognition at every recognition category (the category 1 shown in drawing 1 - category 6).

[0027] Moreover, at the time of recognition, the likelihood of the model corresponding to each category stored in the state-transition model storing memory 7 for recognition by study is computed, and maximum likelihood estimation which makes a recognition result the category corresponding to a model with the maximum likelihood is performed.

[0028] In the dynamic-image recognition approach and dynamic-image recognition search method of a gestalt of operation of this invention, the processing to quantization is the same as that also of the time of recognition at the time of study.
[0029] Hereafter, along with drawing 1 , the gestalt dynamic-image recognition approach and dynamic-image recognition search method of operation of this invention are explained.
[0030] First, the feature-extraction section 2 extracts a DCT multiplier from the MEPG data 1 for retrieval as a feature vector.
[0031] Here, the MEPG data 1 are explained briefly.
[0032] By the MEPG standardization coding method, within a frame, motion compensation vector information was used and data are compressed inter-frame again by the DCT (discrete cosine transform; Discrete Cosine Transform) multiplier for every 8x8-pixel block, and quantization.
[0033] Moreover, each frame of the usual MEPG data 1 consists of coded data of one, I picture, P picture, and B picture, type of three kinds.
[0034] In addition, in coding in a frame, and P picture, forward direction inter-frame predicting coding and B picture mean [I picture] bidirectional inter-frame predicting coding.
[0035] In the usual sequence, one GOP (Group of Picture) starts in I picture, and arranges P picture or B picture at suitable spacing according to violence, demand image quality, etc. of a motion of an image.
[0036] With the gestalt of operation of this invention, in order to use a DCT multiplier, it is used, changing all frames into the image data which is I picture.
[0037] In addition, the conversion to I picture from the MEPG data 1 which consist of I, P, and a B picture is possible by carrying out the direct control of the coded data as indicated by for example, the following reference (b).
[0038] (b) Shin-Fu Chang and David G.Messerchmitt: "A New Approach to Decoding and Compositing Motion-Compensated DCT-Based Images" and Proceedings oflCASSP'93(1993). drawing 2 are drawings showing the outline configuration of the DCT multiplier of the MEPG data 1 and the MEPG data 1.
[0039] As shown in drawing 2 , by the MEPG data 1, the image data of one frame is divided into the MxN block with which 1 block consists of 8x8 pixels, a DCT operation is carried out to the block unit, and, thereby, the DCT multiplier shown in the figures 1-64 within the block of the bottom of drawing 2 is obtained.
[0040] With the gestalt of operation of this invention, a suitable number is taken out for the DCT multiplier (DCT multiplier of the field of E1 shown in drawing 3) of a low-frequency component among the DCT multipliers of this 8x8-pixel block, and this is performed to a whole block and let the numerical train which put the taken-out whole DCT multiplier in order be the feature vector (f) of that frame.
[0041] Since there are 16 blocks in all supposing it uses a 32 pixel x32 pixel image and takes out the DCT multiplier of i pieces from each block, the dimension of the feature vector in this case is set to 16i.
[0042] Since one feature vector (f) is obtained from the image data of one frame of the MPEG data 1, a feature-vector train (F) is acquired from the image data of the continuous frame (screen) which displays a series of dynamic images, and this feature-vector train (F) is recorded on the memory 3 for the description storing.
[0043] In addition, the DCT multiplier used as a feature vector (f) The DCT multiplier on the 1st Rhine horizontal in addition to a number with a suitable low-frequency component of DCT multipliers (DCT multiplier of the field of E2 shown in drawing 3),

You may make it use the DCT multiplier 1st on [of the perpendicular approach] Rhine (DCT multiplier of the field of E3 shown in drawing 3), or the DCT multiplier on the diagonal line containing a dc component (DCT multiplier of the field of E4 shown in drawing 3).
[0044] By using the DCT multiplier on the 1st horizontal Rhine (DCT multiplier of the field of E2 shown in drawing 3) as a feature vector, when the motion with the specific mainly horizontal pattern of a dynamic image is dominant, it is possible to extract the description of a dynamic image with a sufficient precision by the small DCT multiplier.
[0045] Moreover, the specific pattern of a dynamic image is able to extract the description of a dynamic image with a sufficient precision by the small DCT multiplier, when a motion of a perpendicular direction is mainly dominant by using the DCT multiplier 1st on [of the perpendicular approach] Rhine (DCT multiplier of the field of E3 shown in drawing 3) as a feature vector.
[0046] moreover, the thing for which the DCT multiplier on the diagonal line containing a dc component (DCT multiplier of the field of E4 shown in drawing 3) is used as a feature vector -- the specific pattern of a dynamic image -- water square -- when both motions of a method and a perpendicular direction are included, it is possible to extract the description of a dynamic image with a sufficient precision by the small DCT multiplier.
[0047] Furthermore, as a feature vector (f), it also becomes it is possible and possible [that this extracts the description of a dynamic image in a detail more] to use together a DCT multiplier and a motion compensation vector.
[0048] In the quantization section 4, this feature-vector train (F) is changed into a symbol train (O) by vector quantization, and is recorded on the symbol storing memory 5.
[0049] That is, each feature vector is changed into the symbol corresponding to a representation point vector with the nearest distance based on the list of the representation points for the quantization prepared beforehand.
[0050] A code book, and a call and this code book created this representation point group with the LBG algorithm indicated by the following reference (Ha) using a part of feature vector extracted from the image of various kinds of operation.
[0051] (c) Y.Linde, A.Buzo, R.M.Gray; "An Algorithm for Vector Quantizer design", and IEEE Trans.Commin.vol.COM-28(1980). -- k-mean (k-average) indicated in addition by creation of this code book at following reference (**) You may create with an algorithm.
[0052] (d) Supposing it expresses a code book like following the (1) type X.D.Huang, Y.Ariki, M.A.Jack; "Hidden Markov Model for Speech Recognition", and now [Edinburg Univ.Press(1990).], a feature vector (f) will be changed into the symbol (Ot) shown in the following formula (2).
[0053]
[Equation 1]
C=c1, c2, and ..... cN ..... (1)
[0054]
[Equation 2] Ot=vk ..... (2)
k=argminjd(f,cj)
however, d (x y) -- the distance of x and y -- a feature-vector train (F) is changed into a symbol train (O) by the processing so far, and a state-transition model performs study and recognition for this symbol train (O).
[0055] In addition, about the actuation so far, the time of study is the same at the time

of recognition.
[0056] it is indicated by said reference (**) or following reference (**) as this state-transition model -- it hides and the Markov (HMM is called hereafter.) model is used.
[0057] (e) Seiichi Nakagawa; "the speech recognition by the probability model", the Institute of Electronics, Information and Communication Engineers (1990)
At the time of study, the parameter of said HMM model is presumed, and only the number of categories to recognize is prepared at the time of recognition, and the probability for the feature-vector train for recognition (F) to be generated is computed by the likelihood calculation section 8 at it from each of the HMM model stored in the state-transition model storing memory 7 for recognition.
[0058] Hereafter, a HMM model is explained briefly.
[0059] A HMM model is a probable state-transition model, and can be regarded as modeling of the generation source of a time series phenomenon.
[0060] Drawing 4 is the conceptual diagram showing the concept of a HMM model.
[0061] As shown in drawing 4 , two or more conditions (q1-q5) exist in a HMM model, and the probability (aij) which changes from each condition (q1-q5) to other conditions is given to it.
[0062] Along with **, a state transition occurs [time of day] probable, and a symbol (O1-Ot) is further outputted probable from each condition.
[0063] What can be observed is this output symbol train (O=O1, O2, ..., Ot), and cannot carry out direct observation of the condition.
[0064] this -- " -- hiding -- " -- it is the origin of a Markov model.
[0065] In application to recognition of operation, each posture in which it can set working hits a condition, therefore the number of conditions needs to choose a suitable number according to the die length and complexity of the actuation for recognition.
[0066] Moreover, in application to recognition of operation, a symbol output probability can interpret the time series pattern itself and change of telescopic motion etc. of posture change of state transition probability as hitting the part which describes fluctuation of each posture, and fluctuation of the observation of a posture.
[0067] A HMM model is described by the following parameters.
[0068]
[Equation 3] S= {st}: The set of a condition. st t-th condition (it cannot observe)
O=O1, O2, ..., OT ; Observed symbol sequence (die-length T)
A={a ij|a ij=Pr () [second t+1] = j|s t =i}: State transition probability a ij is probability
B= {bj(Ot) |bj(Ot) =Pr (Ot|st=j)} which changes from a condition (si) to a condition (sj). : The symbol output probability bj (k) is set in the condition (sj). Probability pi= which outputs a symbol (upsilonk) {pi i|pi i=Pr (s1=i)} : Study of an initial-state probability, next the time series pattern (symbol train (O)) which used the HMM model, and the procedure of recognition are explained.
[0069] << -- procedure>> at the time of study -- the model parameter estimation section 6 presumes the parameter of a state-transition model which generates the symbol train (O) to the symbol train (O) acquired from the data for study given for every category, and stores it in the state-transition model storing memory 7 for recognition. [two or more]
[0070] The recognition system by the HMM model consists of one HMM model for every category.
[0071] If the HMM model for every category for recognition is now set to lambdai (={Ai, Bi, pii}), study of this lambdai will be performed using the training pattern for every category.

[0072] Here, study is exactly presuming the parameter Ai of a HMM model which is easy to generate a training pattern, i.e., state transition probability, the symbol output probability Bi, and initial-state probability pii.
[0073] In order to presume the parameter of a HMM model from a training pattern, the Baun-Welch algorithm indicated by said reference (d) or reference (e) is used.
[0074] It is the procedure which specifically repeats asking for a model parameter with nearby likelihood higher than it based on the parameter of the procedure which repeats it until it can consider that asking for the parameter of a HMM model with more high likelihood converged enough from the value of likelihood, change, etc., i.e., a certain HMM model, sequentially from a certain initial value.
[0075] The check of convergence is possible by checking the value of likelihood with the forward algorithm indicated by said reference (d) for every repeat.
[0076] It is [0077] when it expresses with a formula.
[Equation 4]
モデル$\lambda = (\pi_i, a_{ij}, b_i(v))$から，より良いモデル$\bar{\lambda} = (\bar{\pi}_i, \bar{a}_{ij}, \bar{b}_i(v))$を，次のようにして求める。

[0078]
[Equation 5]
$$\bar{a}_{ij} = \frac{\sum_{t=1}^{T-1} \xi_t(i,j)}{\sum_{t=1}^{T-1} \gamma_t(i)}; \qquad \cdots\cdots (3)$$

[0079]
[Equation 6]
$$\bar{b}_i(k) = \frac{\sum_{t\in\{t|O_t=v_k\}} \gamma_t(i)}{\sum_{t=1} T \gamma_t(i)}; \qquad \cdots\cdots (4)$$

[0080]
[Equation 7]
$$\bar{\pi}_i = \gamma_1(i). \qquad \cdots\cdots (5)$$

[0081] However, it is here and is [0082].
[Equation 8]
$$\begin{aligned}\gamma_t(i) &\equiv P(s_t = i|O_1, O_2, \ldots, O_T, \lambda) \\ &= \frac{P(s_t = i, O_1, O_2, \ldots, O_T|\lambda)}{P(O_1, O_2, \ldots, O_T|\lambda)} \\ &= \frac{P(O_1, O_2, \ldots, O_t, s_t = i|\lambda)P(s_t = i, O_{t+1}, O_{t+2}, \ldots, O_T|\lambda)}{P(O_1, O_2, \ldots, O_T|\lambda)} \\ &= \frac{\alpha_t(i)\beta_t(i)}{P(O|\lambda)}. \qquad \cdots\cdots (6)\end{aligned}$$

[0083]
[Equation 9]

$$\begin{aligned}
\xi_t(i,j) &\equiv P(s_t = i, s_{t+1} = j | O_1, O_2, \ldots, O_T, \lambda) \\
&= \frac{P(s_t = i, s_{t+1} = j, O_1, O_2, \ldots, O_T | \lambda)}{P(O_1, O_2, \ldots, O_T | \lambda)} \\
&= \frac{P(O_1, O_2, \ldots, O_t, s_t = i | \lambda) a_{ij} b_j(O_{t+1}) P(O_{t+2}, O_{t+3}, \ldots, O_T | \lambda)}{P(O_1, O_2, \ldots, O_T | \lambda)} \\
&= \frac{\alpha_t(i) a_{ij} b_j(O_{t+1}) \beta_{t+1}(j)}{P(O|\lambda)}. \qquad \cdots\cdots (7)
\end{aligned}$$

[0084] The place which said each formula means, (3) types are reevaluation of aij under the HMM model lambda, and (4) types are reevaluation of bi (k) under the HMM model lambda.
[0085] It can ask for the parameter of the state-transition model for recognition corresponding to study data in the above mentioned procedure.
[0086] In this way, the model for every category for which it asked is used in the case of recognition.
[0087] << -- procedure>> at the time of recognition -- the procedure of recognition -- every -- it is carried out by likelihood count of a HMM model, and selection of maximum.
[0088] lambdai calculates the probability (likelihood) (O|lambdai) Pr which outputs the symbol train (O=O1, O2, ..., Ot) which is a recognition object pattern to the pattern for recognition.
[0089] With the forward algorithm indicated by said reference (d), recursively, count of likelihood is the following, and it can make and ask for it.
[0090] That is, the probability Pr (O|lambdai) for a certain model lambda= {A, B, pi} to output a symbol sequence (O=O1, O2, ..., Ot) is [0091].
[Equation 10]

$$\Pr(O|\lambda) = \sum_{i \in S_F} \alpha_T(i) \qquad \cdots\cdots (8)$$

[0092] However, SF is the set of a final state and alphaT (i) is [0093] here.
[Equation 11]

$$\alpha_t(i) \equiv \Pr(O_1, O_2, \ldots, O_t, s_t = i | \lambda) \qquad \cdots\cdots (9)$$

[0094] In the value come out of and defined, the HMM model lambda generates a symbol sequence (O=O1, O2, ..., Ot), and it is the probability which is a condition (St=i) in time amount t.
[0095] This is [0096].
[Equation 12]

$$\alpha_t(j) = \{\sum_i \alpha_{t-1}(i) a_{ij}\} b_j(O_t) \qquad \cdots\cdots (10)$$

$$\alpha_1 = \pi_i b_i(O_1) \qquad \cdots\cdots (11)$$

[0097] It asks by ********.
[0098] In this way, the category (Gk) to lambdai of likelihood max and (k=argmaxiPr (O|lambdai)) are chosen as a recognition result, and are stored in the memory 6 for recognition results from the model with which the called-for likelihood serves as max, i.e., Pr calculated by the formula (11) from the formula (1), (O|lambdai).
[0099] Moreover, at the time of retrieval, it searches [whether the part of the MEPG

data 1 throat used as the candidate for retrieval serves as likelihood max to the HMM model corresponding to the candidate for retrieval, and] by scanning the inside of the MEPG data 1.
[0100] In this case, in order to ask for the maximum likelihood part in the MEPG data 1 efficiently, it is possible to use the HMM spotting algorithm indicated by said reference (e).
[0101] Recognition by the HMM model is performed by maximum likelihood estimation, and study is realized in the form of presumption of the parameter of the HMM model from the data for study so that clearly from the above processing flow.
[0102] And since likelihood count is performed from the whole symbol sequence, if the symbol train pattern peculiar to a category has appeared, there is a merit of being strong, to migration of some of directions of a time-axis, telescopic motion, etc.
[0103] Moreover, it asks for each [of the time series pattern of a dynamic image] likelihood of a time, and retrieval of a specific time series pattern is attained by performing threshold processing etc. to this.
[0104] Next, two person actuation check experimental results for a tennis actuation image are explained as an example of an experimental result based on the gestalt of operation of this invention.
[0105] [Experiment 1] In the gestalt of operation of this invention, an example of the photograph of the tennis actuation image used for the experiment 1 is shown in drawing 5 .
[0106] from the tennis actuation image shown in the upper case of drawing 5 , it is shown in the lower berth of drawing 5 -- as -- a background -- the MEPG data created based on the example of an image from which difference extracted the person field and this person field was extracted -- the candidate for recognition -- carrying out -- DCT -- the recognition engine performance when making counting into characteristic quantity was evaluated.
[0107] the recognition engine performance -- the DCT multiplier of per each block (8x8 pixels) -- low -- the order from degree component -- one every train, 1, 3, 6, 10, and 15, -- 21 or 28 pieces were extracted, it experimented, respectively, and the recognition rate was searched for. [i.e.,]
[0108] In addition, it becomes only DC component when the DCT multiplier per each block is 1.
[0109] Moreover, image size was made into two kinds, 16x16 pixels (it is 1x1 block at a macro block unit), and 32x32 pixels (it is 2x2 blocks at a macro block unit).
[0110] Moreover, size of the code book for quantization is set to 48 in each class size 8 and 6 class sum total, and it creates with a LBG algorithm, and the number of conditions of a HMM model is 12, and the number of symbols is 48.
[0111] Drawing 6 is a photograph in which the target tennis actuation image is shown in the experiment 1 of the gestalt of operation of this invention.
[0112] As shown in drawing 6 , the target tennis actuation is six categories of a backhand oystershell (back-volley), a backhand stroke (back-stroke), a forehand oystershell (fore-volley), a forehand stroke (fore-stroke), a smash (smash), and service (service).
[0113] About each of actuation of six categories, the image data of 10 trial of operation was collected, five of the trial of this were used as data for study, the parameter of a HMM model was presumed, and the recognition experiment was conducted by using five remaining trial as a test data.
[0114] In this case, it experimented by changing into ten kinds the selection approach

which chooses five trial from from among 10 trial.
[0115] Therefore, a recognition rate is estimated by how many times of the recognition experiments of 5x10x6=300 time it succeeded.
[0116] This recognition experimental result is shown in Table 1 and Table 2.
[0117]
[Table 1]
表 1: 認識実験結果 (16x16)

DCT(per Block)	認識率 (%)
1	72.66
3	93.00
6	97.66
10	99.33
15	96.00
21	98.00
28	96.33

[0118]
[Table 2]
表 2: 認識実験結果 (32x32)

DCT(per Block)	認識率 (%)
1	88.33
3	88.66
6	98.00
10	98.00
15	99.66
21	100.00
28	99.66

[0119] The recognition rate was improving greatly and by increasing the DCT multiplier used as characteristic quantity showed that the DCT multiplier of a low-frequency component was comparatively effective as characteristic quantity for the image recognition of person actuation so that he could understand from Table 1 and display 2.
[0120] Moreover, even when an object image was comparatively small, by using a DCT multiplier to a high frequency component showed that 98% or more of recognition rate was acquired, and a recognition rate without the case where an image is large, and inferiority could be realized.
[0121] [Experiment 2] In the gestalt of operation of this invention, the application experiment to dynamic-image retrieval was conducted for a series of dynamic-image data including two or more sorts of actuation.
[0122] The learned HMM model of each category of operation examined whether actuation could be searched by choosing the HMM model of likelihood max in each time.
[0123] The screen size used 32x32 pixels, and the DCT multiplier was considered as per [6] each block as characteristic quantity.
[0124] Drawing 7 is a graph in the gestalt of operation of this invention which shows the experimental result of experiment 2.
[0125] drawing 7 -- each observation of a time -- being based -- respectively -- the logarithm of the HMM model of six categories -- it is the graph which plotted

likelihood.
[0126] Therefore, it is expected that likelihood will serve as max at the time of termination of operation.
[0127] From the graph shown in drawing 7 , he can check that the HMM model of each set elephant actuation is order with maximum likelihood, and can understand that logging of the section of operation is possible by threshold processing.
[0128] Thereby, the retrieval of a specific pattern of operation in continuation dynamic-image data is possible.
[0129] In addition, in explanation of the gestalt of operation of said this invention, although the MEPG data encoded by the standard coding method of MEPG and MEPG2 grade were used, it cannot be overemphasized that the data which are not limited to this and encoded by standard coding methods, such as motion-JPEG, can be used.
[0130] As mentioned above, although this invention was concretely explained based on the gestalt of implementation of invention, it cannot be overemphasized that it can change variously in the range which this invention is not limited to the gestalt of implementation of said invention, and does not deviate from the summary.
[0131]
[Effect of the Invention] It will be as follows if the effectiveness acquired by the typical thing among invention indicated by this application is explained briefly.
[0132] (1) According to this invention, since the DCT multiplier or the DCT multiplier, and the motion compensation vector were used as characteristic quantity, it becomes possible to direct-recognize and to search a specific dynamic-image pattern from the dynamic-image data of the few capacity compressed by the standard coding method of MEPG and MEPG2 grade.
[0133] This becomes possible to lessen the processing time of data processing.
[0134] (2) Even if there are migration of some of directions of a time-axis, telescopic motion, etc. if the feature-vector train pattern peculiar to a category has appeared since likelihood count is performed from the whole feature-vector sequence according to this invention, it becomes possible to recognize and search a specific dynamic-image pattern with a sufficient precision.
[0135] (3) According to this invention, it is possible to raise a recognition rate by being able to raise a recognition rate sharply by using the DCT multiplier used as characteristic quantity to a high frequency component, and using the DCT multiplier used as characteristic quantity to a high frequency component, even if it is the case that an object image is comparatively small.
[0136] (4) According to this invention, it is widely applicable to logging for a right hand side of a request etc. from animations, such as a doubtful action monitor in a bank or a store, and a sport.

[Translation done.]

DESCRIPTION OF DRAWINGS

[Brief Description of the Drawings]
[Drawing 1] It is the functional block diagram showing the outline configuration of the dynamic-image recognition retrieval equipment with which the dynamic-image recognition approach and dynamic-image recognition search method which are the gestalt of implementation of 1 invention of this invention are applied.
[Drawing 2] It is drawing showing the outline configuration of the DCT multiplier of the MEPG data 1 and the MEPG data 1.
[Drawing 3] It is drawing for explaining the extract approach of a DCT multiplier in the gestalt gestalt of operation of this invention.
[Drawing 4] It is the conceptual diagram showing the concept of a HMM model (hiding Markov).
[Drawing 5] In the gestalt of operation of this invention, it is the halftone image displayed on the display in which the example of the tennis actuation image used for the experiment 1 is shown.
[Drawing 6] It is the halftone image displayed on the display in which the target tennis actuation image is shown in the experiment 1 of the gestalt of operation of this invention.
[Drawing 7] It is the graph which shows the experimental result of the experiment 2 of the gestalt of operation of this invention.
[Description of Notations]
2 [-- The memory for symbol train storing, 6 / -- The model parameter estimation section, 7 / -- The state-transition model storing memory for recognition, 8 / -- The likelihood calculation section, 9 / -- Memory for recognition results.] -- The feature-extraction section, 3 -- The memory for the description storing, 4 -- The quantization section, 5

[Translation done.]

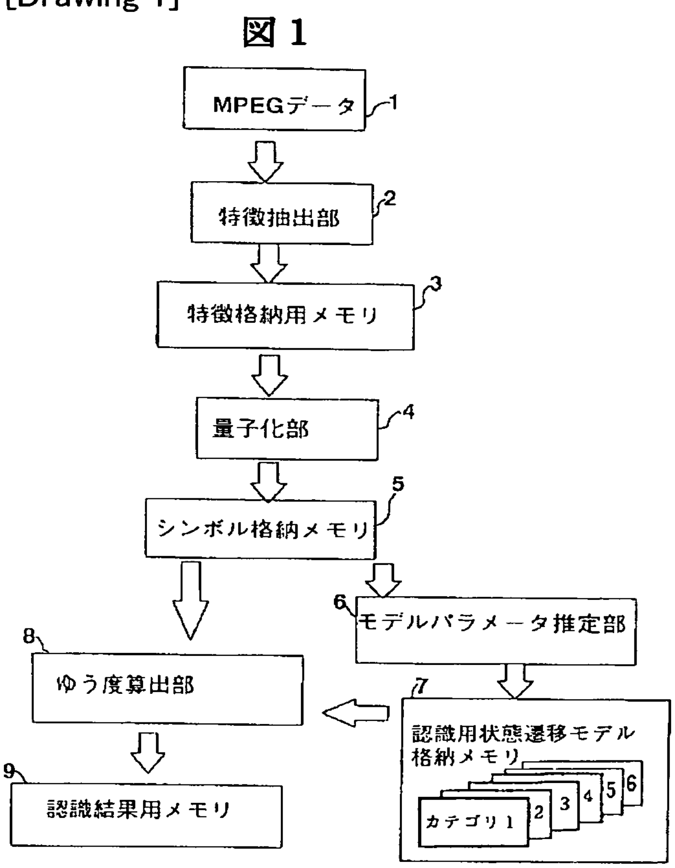
* NOTICES *

JPO and NCIPI are not responsible for any
damages caused by the use of this translation.

1.This document has been translated by computer. So the translation may not reflect the original precisely.
2.**** shows the word which can not be translated.
3.In the drawings, any words are not translated.

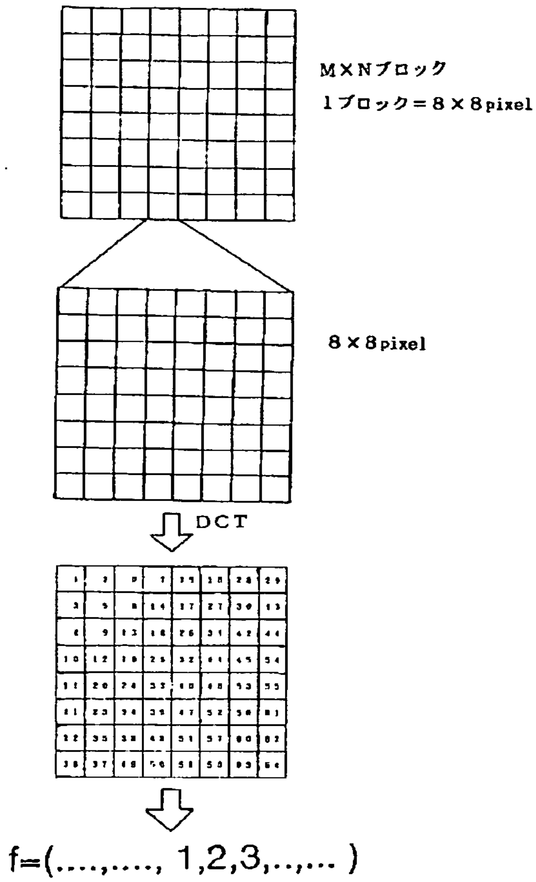
DRAWINGS

[Drawing 1]



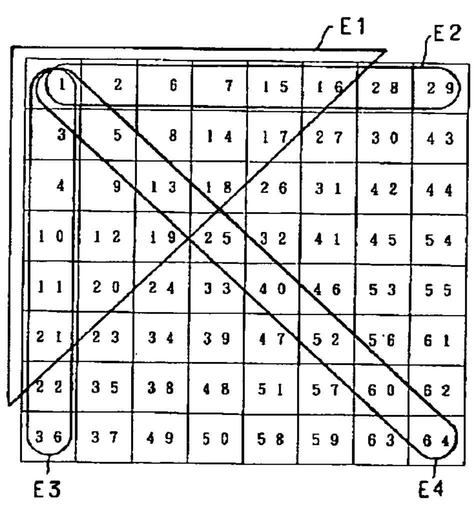
[Drawing 2]

図 2



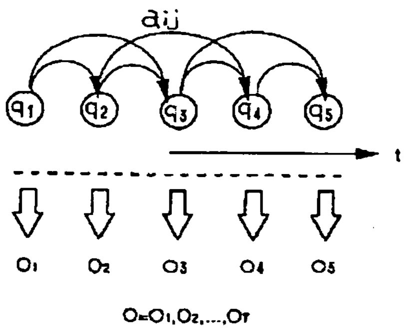
[Drawing 3]

図 3



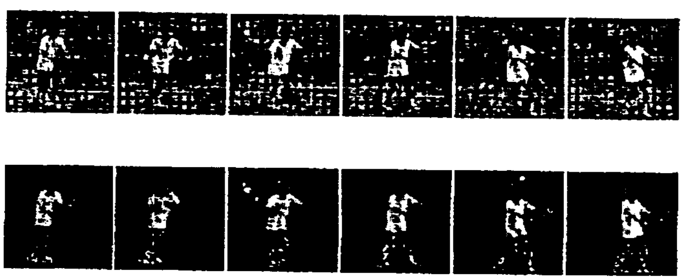
[Drawing 4]

図 4



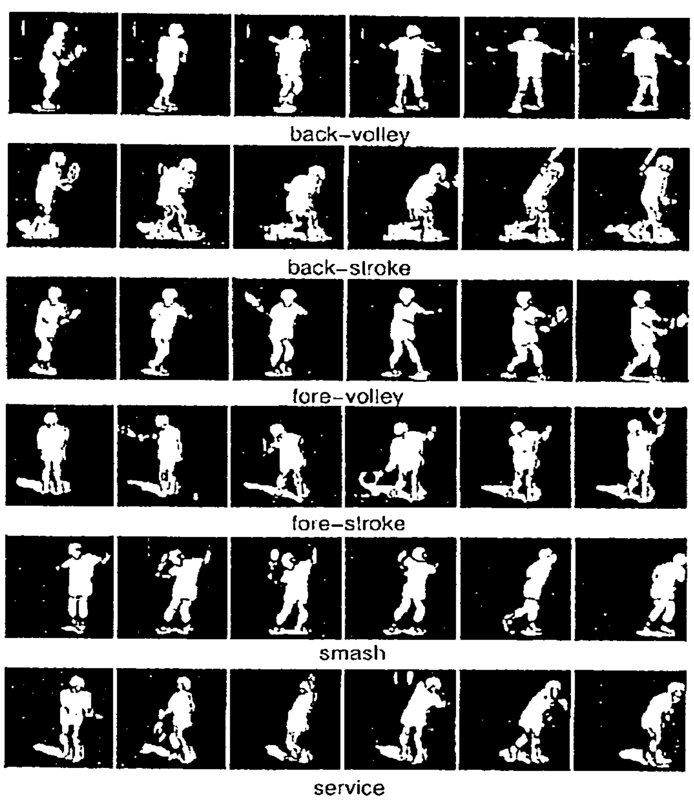
[Drawing 5]

図 5



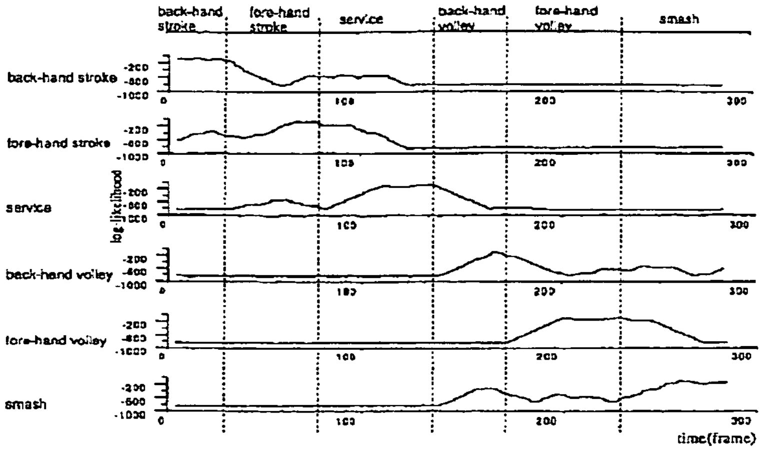
[Drawing 6]

図 6



[Drawing 7]

図 7



[Translation done.]

WRITTEN AMENDMENT

--- [a procedure revision]
[Filing Date] June 27, Heisei 8
[Procedure amendment 1]
[Document to be Amended] Specification
[Item(s) to be Amended] Claim
[Method of Amendment] Modification
[Proposed Amendment]
[Claim(s)]
[Claim 1] In the dynamic-image recognition approach of recognizing the dynamic-image pattern of a series of dynamic images The step which extracts a break and the DCT multiplier of each block for the image data of each screen which displays a series of dynamic images to the block of MxN, By the feature-vector train of the step which extracts at least one of said the DCT multipliers as a feature vector of each screen, and the time series which consists of feature vectors of each screen which displays a specific dynamic-image pattern The step which learns a probable state-transition model for every dynamic-image pattern of two or more specification with which it becomes a recognition key, The feature-vector train of the time series which consists of feature vectors extracted from the image data of each screen which displays a series of dynamic images which are the candidates for recognition, The dynamic-image recognition approach characterized by providing the step which outputs the dynamic-image pattern of the state-transition model with which the likelihood to two or more state-transition models obtained by said study serves as max as a recognition result.
[Claim 2] The dynamic-image recognition approach indicated by claim 1 characterized by the image data of each screen which displays a series of dynamic images which are said candidates for recognition using a part of DCT multiplier contained in the image data of each screen which is compressed by the standard coding method and compressed by the standard coding method as a feature vector of each screen.
[Claim 3] The dynamic-image recognition approach indicated by claim 1 characterized by using a motion vector with a DCT multiplier as a feature vector of each of said screen.
[Claim 4] The dynamic-image recognition approach indicated by claim 3 to which it is characterized by using the part and motion compensation vector of the DCT multiplier contained in the image data of each screen which the image data of each screen which displays a series of dynamic images which are said candidates for recognition is compressed by the standard coding method, and was compressed by the standard coding method as a feature vector of each screen.

[Claim 5] The dynamic-image recognition approach indicated by claim 2 or claim 4 characterized by using it among the DCT multipliers contained in the image data of each screen compressed by said standard coding method, carrying out the feature vector of 3 thru/or the DCT multiplier of 21 low-frequency components.
[Claim 6] The dynamic-image recognition approach indicated by claim 2 or claim 4 characterized by using it among the DCT multipliers contained in the image data of each screen compressed by said standard coding method, carrying out the feature vector of the DCT multiplier on the 1st horizontal Rhine.
[Claim 7] The dynamic-image recognition approach indicated by claim 2 or claim 4 characterized by using it among the DCT multipliers contained in the image data of each screen compressed by said standard coding method, carrying out the feature vector of the DCT multiplier 1st on [of the perpendicular approach] Rhine.
[Claim 8] The dynamic-image recognition approach indicated by claim 2 or claim 4 characterized by using it, carrying out the feature vector of the DCT multiplier on the diagonal line which contains a dc component among the DCT multipliers contained in the image data of each screen compressed by said standard coding method.
[Claim 9] In the dynamic-image recognition search method which extracts the time domain containing a specific dynamic-image pattern out of a series of dynamic images The step which extracts a break and the DCT multiplier of each block for the image data of each screen which displays a series of dynamic images to the block of MxN, By the feature-vector train of the time series which consists of feature vectors of each screen which displays the step which extracts at least one of said the DCT multipliers as a feature vector of each screen, and the specific dynamic-image pattern used as a search key In the feature-vector train of the time series which consists of a step which learns a probable state-transition model, and a feature vector extracted from the image data of each screen which displays a series of dynamic images which are the candidates for retrieval The dynamic-image recognition search method characterized by providing the step which outputs the time domain where the likelihood to the state-transition model obtained by said study is high as a retrieval result.
[Claim 10] The dynamic-image recognition search method indicated by claim 9 characterized by the image data of each screen which displays a series of dynamic images which are said candidates for retrieval using a part of DCT multiplier contained in the image data of each screen which is compressed by the standard coding method and compressed by the standard coding method as a feature vector of each screen.
[Claim 11] The dynamic-image recognition search method indicated by claim 9 characterized by using a motion vector with a DCT multiplier as a feature vector of each of said screen.
[Claim 12] The dynamic-image recognition search method indicated by claim 11 to which it is characterized by using the part and motion compensation vector of the DCT multiplier contained in the image data of each screen which the image data of each screen which displays a series of dynamic images which are said candidates for retrieval is compressed by the standard coding method, and was compressed by the standard coding method as a feature vector of each screen.
[Claim 13] The dynamic-image recognition search method indicated by claim 10 or claim 12 characterized by using it among the DCT multipliers contained in the image data of each screen compressed by said standard coding method, carrying out the feature vector of 3 thru/or the DCT multiplier of 21 low-frequency components.
[Claim 14] The dynamic-image recognition search method indicated by claim 10 or claim 12 characterized by using it among the DCT multipliers contained in the image

data of each screen compressed by said standard coding method, carrying out the feature vector of the DCT multiplier on the 1st horizontal Rhine.

[Claim 15] The dynamic-image recognition search method indicated by claim 10 or claim 12 characterized by using it among the DCT multipliers contained in the image data of each screen compressed by said standard coding method, carrying out the feature vector of the DCT multiplier 1st on [of the perpendicular approach] Rhine.

[Claim 16] The dynamic-image recognition search method indicated by claim 10 or claim 12 characterized by using it, carrying out the feature vector of the DCT multiplier on the diagonal line which contains a dc component among the DCT multipliers contained in the image data of each screen compressed by said standard coding method.

[Translation done.]